Conservation Genet Resour
DOI 10.1007/s12686-014-0246-0

MICROSATELLITE LETTERS

Development and characterization of genomic microsatellite markers in *Rhododendron arboreum*

Shruti Choudhary · Sapna Thakur · Ram Gopal Saini · Pankaj Bhardwaj

Received: 22 May 2014 / Accepted: 2 June 2014

Abstract Population genetics characteristics are the fundamentals of conservation and management practices. *Rhododendron arboreum*, a key biodiversity component inhabiting Indian Himalayas, suffers from overexploitation and global warming. Using biotin–streptavidin hybridization technique, 41 microsatellite markers were designed from an enriched DNA library to provide a genetic background and an insight into the population structure of the species. With a range of 2–14 alleles amplified from 38 loci, the populations were reported with observed and expected heterozygosity of 0.167–0.933 and 0.422–0.917 respectively. Some of the loci showed significant deviations from Hardy–Weinberg equilibrium and overall no linkage disequilibrium was detected. These markers will support genetic diversity and further genotyping studies in *R. arboreum.*

Keywords *Rhododendron arboreum* · Genomic microsatellites · Selective hybridization

Rhododendron arboreum (a member of Ericaceae), a keystone species of the Indian Himalayan region, is an evergreen tree of ecological and medicinal importance. Inhabiting subalpine to alpine transition zone, the Rhododendron forest line supports a regime of biodiversity components. Snow cover favours the survival of biota at higher altitudes during harsh climate. An isothermic shift due to global warming; overexploitation and other anthropogenic activities can pose a threat to the species richness of Himalayan region in near future (Xu et al. 2009). Since no genetic base is available for this species, it becomes prerequisite to unravel the population dynamics essential for planning and adopting conservation measures. Co-dominant microsatellites have emerged as potent molecular markers for population and evolutionary studies. In view of the above context, we developed genomic SSRs to interpret genetic diversity and population structure of the species.

Genomic DNA was isolated from leaves according to the CTAB procedure (Doyle and Doyle 1987). A library enriched for dinucleotide repeats (GA) was constructed using two restriction enzymes (Mse1 and EcoR1) following the protocol standardized by Bhardwaj et al. (2013). The size selected adapter ligated fragments were enriched using biotin–streptavidin capture with $(GA)_{10}$ as probe, amplified and cloned into chemically competent *Escherichia coli* host using pGEM-T EASY vector. Of the 1,200 transformed colonies screened, 351 clones were subsequently sequenced by 3,730 × l DNA Analyzer (Applied Biosystems) with universal M13 as primer. 191 sequences were identified with di- and tri- repeats of lengths ranging from 5 to 23 using SSR Identification Tool and were classified as perfect, compound or interrupted motifs of AG, AC, ACA repeats (or their complements). The sequences were used to design 41 unique, non-redundant primer pairs from Primer 3.

Of the isolated markers, 38 loci were amplified in 10 µl reaction volume, constituting 25 ng template, 0.3 U *Taq* DNA polymerase (Bangalore Genei^{TM}), 1X *Taq* buffer (1 mM Tris pH 9.0, 50 mM KCl 0.01 % gelatin, 1.5 mM $MgCl_2$), 2.5 mM of each dNTPs, 5 ng each of forward and reverse primer. PCR thermal profile comprised of an initial denaturation step at 94 °C for 3 min followed by 35 cycles

Electronic supplementary material The online version of this article (doi:10.1007/s12686-014-0246-0) contains supplementary material, which is available to authorized users.

S. Choudhary · S. Thakur · R. G. Saini · P. Bhardwaj (✉)
Molecular Genetics Laboratory, Centre for Biosciences, Central University of Punjab, Bathinda 151001, India
e-mail: pankajihbt@gmail.com; pankajbhardwaj@cup.ac.in

Published online: 11 June 2014

Table 1 Characteristics of 38 SSR loci developed from *Rhododendron arboreum*

S. no.	Locus name	Primer sequences	Repeat motif	T_a (°C)	N_a	Approx. size range (bp)	Heterozygosity		PIC	Accession no.
							H_o	H_e		
1.	R394	F: GGAAAGTGTGGGTGTTAGTGC R: TTGAGAGATGGCGAGAGAGAG	$(TC)_{16}$	59	4	145–165	0.6667*	0.7181	0.104	KJ851157
2.	RE101	F: GACGGGAATGAGCAAGGTTG R: CTTCAATTCTGCAAGCCCGA	$(AG)_{16}$	55	6	210–240	0.7333*	0.8418	0.480	KJ851180
3.	RF87A	F: TGGGTCATGTTCTGGAAGGT R: TGAACTAACCCTAGCCACACT	$(AG)_{10}(AG)_{10}$ $(GA)_9$	55	6	140–170	0.5677***	0.7689	0.436	KJ851183
4.	RA50	F: ACTCCCTCCTGTCGTTCCTT R: AATCGTGCATCCGTATCCTG	$(TC)_{16}$	58	5	216–234	0.6667	0.7588	0.310	KJ851162
5.	RA324	F: GCGTACAACATGCCCAAATA R: CCCTGTTCTCATTGCTCACA	$(AG)_8(GA)_8$ $(GA)_g$	55	2	184–194	0.4667	0.7994	0.709	KJ851168
6.	RA351	F: TGTCGCTCTCTCACTGATCG R: TTTGTAGTTTTCCCGTGTCCTT	$(AG)_{12}(AG)_{11}$	59	5	338–360	0.5677*	0.7525	0.374	KJ851170
7.	RA346	F: CGGAGCAAGCTCTCTTATCG R: CCTCTCCTGTGTAGCAAGTCG	$(TC)_9$	59	8	100–116	0.5333*	0.7023	0.396	KJ851184
8.	RA321	F: AGAGATGGGTTTGTGTAAAGTCTG R: TATTTCGCTGCCACCCTAAC	$(GA)_9$	55	4	264–280	0.4677*	0.7124	0.373	KJ851185
9.	RA470	F: AGGGACAAGAAGAAGCCACA R: TCGCGCTTATTACAGCTCTTC	$(GA)_{14}(AG)_{10}$ $(AG)_{10}$	55	3	150–154	0.3000**	0.5994	0.438	KJ851173
10.	RF304	F: TCCTAGGGTTTGTTCGCAAT R: TGCTAGCGATTTCCTAGGGT	$(AG)_{13}(GT)_9$ $(AG)_5(GT)_9$	55	5	218–230	0.2333*	0.7339	0.811	KJ851178
11.	RA254	F: AGTAGCAACACCCACACACT R: GGAGGGGCTGTAGTCTGATT	$(CT)_{16}(CT)_{10}$	55	7	150–164	0.8000*	0.8260	0.413	KJ851179
12.	RA337	F: GAGCGAGAGAGAGGTGTTGG R: ATTCACGGGAATCTTCACCA	$(AG)_6$	59	2	206–208	0.2000	0.4316	0.266	KJ851169
13.	RA267	F: ACGGAGAAGCAGTGAGCATT R: TGCACAGGAACACCCAATAA	$(GA)_{11}(AG)_{10}$	59	3	196–200	0.3000**	0.7102	0.510	KJ851163
14.	RA272A	F: GCCCCGGTGACTCATAAAAT R: TGGTACAAGTGGGACACGA	$(CT)_8(CT)_{11}$	59	3	188–194	0.2667*	0.7525	0.683	KJ851164
15.	R460	F: CCCTACTTCTTTCATCACATACAA R: CAACTCCGGTCATTTTTGGT	$(GA)_{13}(AG)_{12}$	59	5	188–196	0.4333	0.7345	0.621	KJ851158
16.	R97	F: AGCAGCAACAATGGTGTCC R: TCTAGAAGGCCTCCCATTCC	$(AG)_{10}(AG)_{13}$	59	2	188–190	0.5000	0.6672	0.381	KJ851160
17.	RA430	F: GCGTAAATCGAGTTCGGAAG R: CTCTCTCTAATCGAATTCCCG	$(TC)_{10}(CT)_6$	59	7	166–180	0.6333	0.8627	0.610	KJ851171

Table 1 continued

S. no.	Locus name	Primer sequences	Repeat motif	T_a (°C)	N_a	Approx. size range (bp)	Heterozygosity		PIC	Accession no.
							H_o	H_e		
18.	R356	F: GAGCTAAGCACGCCGTATTC R: AAATTCGACGGCAAAGAGG	$(TC)_9$	59	4	186–194	0.6667*	0.6774	0.104	KJ851166
19.	RA134	F: GGAGAGAGAGGCCGAGAGAG R: ACGTCGCTCTTGTCAAGCAT	$(GA)_6(GA)_6$ $(AG)_8$	59	3	220–228	0.3000	0.4220	0.140	KJ851167
20.	RF103	F: GATAGAGAGACAGGGGCAGC R: TGTACGCCAAGACTCCCATT	$(GA)_{14}$	59	5	288–300	0.5333	0.8023	0.482	KJ851181
21.	RF29	F: ACAGACAGAAGCAGCGGAAC R: AAGGGGAGGAGATCGAGTTG	$(GA)_{10}(AG)_{13}$	59	9	138–160	0.9333	0.8254	0.266	KJ851174
22.	RF43	F: AATTCGATGGGTTGGTGGTA R: GCCTTCTCTGTTCTCGGTTTT	$(AG)_{10}$	59	4	190–200	0.3667*	0.5294	0.237	KJ851175
23.	RF245	F: GGGTTTTTGATCTTCATACGG R: AATCGGTTCAAGAGGGGTTT	$(AG)_{11}$	55	8	198–220	0.5333	0.8576	0.693	KJ851177
24.	RF87B	F: GGGGAAAGGTCATTGGAGAT R: TTCTGAACTAACCCTAGCCACA	$(GA)_{10}(GA)_{10}$ $(AG)_{10}$	55	3	310–316	0.4000*	0.7249	0.536	KJ851183
25.	RF98	F: AATCCCATCCCCTAACTTGG R: CCGTGGCTTTACCTTTCACT	$(AG)_{22}$	59	10	170–196	0.6667	0.9169	0.720	KJ851182
26.	RA7	F: GTCTACAATGCTTGCTTCCG R: CCTTATTTATTCTCTCTCT	$(AG)_{10}(AG)_9$	55	5	110–120	0.7333*	0.7881	0.370	KJ851150
27.	RA19	F: AGCCAAAAACTTTCTTCTTTCC R: CTGTCGGCTGTCAGAGTTGA	$(GA)_{19}$	55	4	124–138	0.5333	0.7073	0.381	KJ851151
28.	RF71	F: GCGTACAACATGCCCAAATA R: GTCGTTGCAGTTCAATCTCG	$(AG)_{12}$	55	5	300–312	0.4667	0.8147	0.710	KJ851176
29.	RA443	F: CCATGCCTGAAGCAAACAC R: AGACTCCAAAGTCCTATCTGTGC	$(AG)_{12}$	59	8	184–200	0.4000*	0.8282	0.649	KJ851162
30.	RA272B	F: ATGCAATGGAAATGGGAAAG R: GGAACGGGTAATTCGGATCT	$(CT)_8(CT)_{11}$	59	3	184–194	0.3333	0.5085	0.194	KJ851164
31.	R372	F: GGTTGGGTGGATGGAGTAAC R: GCAATTTGCATAGCACTGTAAT	$(AG)_{15}(AG)_{15}$	55	5	230–244	0.4667	0.8237	0.600	KJ851156
32.	R422	F: GCGGTACTGTTCCGATCAC R: TCCCAGCTCATCCACACATA	$(AG)_{12}(GA)_{13}$	55	12	146–172	0.7333	0.8593	0.624	KJ851155
33.	R304	F: TCCTAGGGTTTGTTCGCAAT R: GCGTATTGTTGCACGAAAAA	$(AC)_9(CT)_{12}(TC)_5$ $(AC)_9(TC)_{13}$	55	4	380–410	0.3000**	0.6582	0.482	KJ851161
34.	R79	F: AACGTGAAAACTGAGCAGAAC R: CCCGTGGGTAGGAAAATCAT	$(AG)_{10}(AG)_5(AG)_{10}(AG)_7$	55	4	236–250	0.2667	0.5316	0.358	KJ851159

Table 1 continued

S. no.	Locus name	Primer sequences	Repeat motif	T_a (°C)	N_a	Approx. size range (bp)	Heterozygosity		PIC	Accession no.
							H_o	H_e		
35.	RA276	F: GCCAAAAAGCATCAAGTCGT R: TGTGATTTTTGTGTTGGATGG	$(TC)_{11}$	55	6	390–420	0.4000*	0.6605	0.421	KJ851165
36.	R25	F: CCAACAACCCGAGAAAAAGA R: AGTGGGTTTCCGAGACAAAG	$(AG)_{12}(GA)_{11}$	55	14	164–200	0.7667	0.9119	0.709	KJ851154
37.	RA30	F: TGTGATTTTTGTGTAGGATGGT R: AGCAACCCACTTCTCCTCTTC	$(AG)_{17}$	55	2	150–160	0.1667*	0.5814	0.651	KJ851152
38.	RA54	F: CAAAAATGGCCAACAAGGAT R: ATTGCCTCCATACACAACCA	$(CT)_5(CT)_6(CT)_{17}$	55	2	190–200	0.5000	0.6605	0.429	KJ851153

T_a annealing temperature, N_a total number of alleles, H_o observed heterozygosity, H_e expected heterozygosity, *PIC* polymorphic information content

Significant deviations from Hardy–Weinberg equilibrium at * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

of denaturation at 94 °C for 1 min, annealing at specific temperature (standardized for each locus as depicted in Table 1) for 1 min and elongation at 72 °C for 1 min; and a final elongation step at 72 °C for 8 min. These loci were characterized among three populations of *R. arboreum* (each comprising of 10 individuals) from Dalhousie, Ghatasni, Multhan (India; geographical coordinates given in Supplementary Table 2) on 6 % denaturing polyacrylamide gel. Number of alleles ranged from 2 to 14 (an average of 5.2) with fragment length of range 100–420 bp. An overall population structure analysis was performed using PopGene version 1.31 (Yeh et al. 1999). Polymorphic information content and observed and expected heterozygosity varied within the range of 0.104–0.911 (with an average of 0.464), 0.167–0.933 (with a mean of 0.523) and 0.422–0.917 (with an average of 0.723) respectively. 19 loci reported significant deviations from Hardy–Weinberg equilibrium (Table 1) which might be due to natural selection or other factors. As a whole, no significant linkage disequilibrium was detected at the population level ($p < 0.05$). Also, the population-wise diversity parameters (Supplementary Table 1) were estimated using GenAlEx version 6.5 (Peakall and Smouse 2012). The polymorphic loci defined here are highly informative and will be used further in genotyping or population genetics studies for managing conservation policies for *R. arboreum* and also, for predicting gene flow and profound effects of other evolutionary forces disturbing HWE.

Acknowledgments The research work present in the manuscript is funded by Department of Science and Technology, Govt. of India. This is CUPB communication no P78/14.

References

Bhardwaj P, Kumar R, Sharma H, Tewari R, Ahuja PS, Sharma RK (2013) Development and utilization of genomic and genic microsatellite markers in Assam tea (*Camellia assamica* ssp. *assamica*) and related *Camellia* species. Plant Breed 132(6): 748–763

Doyle J, Doyle J (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel: population genetic software for teaching and research-update. Bioinformatics 28(19):2537–2539

Xu J, Grumbine RE, Shrestha A, Eriksson M, Yang X, Wang Y, Wilkes A (2009) Melting Himalayas: cascading effects of climate change on water, biodiversity, and livelihoods. Conserv Biol 23(3):520–530

Yeh F, Yang RC, Boyle T (1999) PopGene Version 1.31: Microsoft window based freeware for population genetic analysis. University of Alberta and Centre for International Forestry Research 11–23